Published in final edited form as:

Learn Individ Differ. 2016 February ; 46: 11–16. doi:10.1016/j.lindif.2014.11.024.

Getting Something out of Nothing: Analyzing Patterns of Null Responses to Improve Data Collection Methods in sub-Saharan Africa

Sascha Hein[1,*,+], **Jodi Reich**[1,2,+], **Sarah Marks**[3], **Philip E. Thuma**[4], and **Elena L. Grigorenko**[1,5]

[1]Yale University, USA

[2]Temple University, USA

[3]University of North Carolina at Chapel Hill, USA

[4]Macha Research Trust, Zambia

[5]Moscow State University of Psychology and Education, Russia

Abstract

Careful development and adaptation of assessments is imperative for cultural psychological research. However, despite the best efforts, the use of assessments in new contexts can reveal atypical and/or unexpected patterns of performance. We found this to be the case in the testing of assessments to be used for a larger investigation of Specific Reading Disabilities in Zambia. In a sample of 207 children (100 female) from grades 2 to 7, we illustrated that assessment characteristics (i.e., stimulus type, answer choice, and response type) differentially impact patterns of responsiveness. The number of missing values was highest for assessments that (1) used written stimuli, (2) had an open-ended answer choice, and (3) required an action response. Age and socio-economic status explained some of the variance in responsiveness in selected, but not all assessments. Consideration of the impact of stimulus and response types when adapting assessments cross-linguistically and cross-culturally is essential.

Keywords

assessment design; assessment adaptation; assessment translation; sub-Saharan Africa; Zambia; missing data

Worldwide, assessments for research in less extensively studied countries, cultures, and languages are not readily available. Therefore, research in a country, such as Zambia, the

[*]Corresponding Author: Sascha Hein, Child Study Center, Yale University, PO Box 207900, 230 South Frontage Road, New Haven, CT 06520-7900, USA, Sascha.Hein@yale.edu, Phone: (203) 785-4831.

[+]Both authors contributed equally to this manuscript

location of the research project presented here, is often faced with the challenge of developing and adapting assessments required for the particular purpose of the study. The adaptation of reliable, linguistically valid and bias-free assessments are major tasks for cultural psychology (Peña, 2007; Van de Vijver & Hambleton, 1996). Cross-cultural literature has emphasized the importance of establishing linguistic, cultural, functional, and metric equivalence for assessments in new contexts (Hambleton & Kanjee, 1995; Peña, 2007). Even if these expectations are met, empirical examination of whether the utilized assessment procedures are adequate for new applications is necessary (Van de Vijver & Hambleton, 1996). Adequate test and assessment development is a growing concern worldwide (International Test Commission, 2006); however, even when assessments are linguistically appropriate and were properly adapted for the specific context in which they are used, data collection can reveal challenges. In this brief report, we aim to illustrate the relation between null responses on a variety of reading-related assessments and characteristics pertaining to the assessment and the examinee.

Three major factors have been suggested as sources for biases in inferring examinees‘ performance on assessments: construct, method, and item (e.g., van de Vijver & Poortinga, 1997; van de Vijver & Poortinga, 2005). First, the construct being assessed may have a different definition and notion across cultures, with potential differential appropriateness of item and stimulus content. Second, the assessment method could induce response biases, which could be due to the assessment itself (e.g., children‘s lack of familiarity with the response procedures and the stimuli, complex instructions) or to its administration (e.g., assessment conditions and assessor effects). Third, features of items such as poor wording or inadequate coverage of the curriculum can lead to bias in the assessment. Here, we report on an observation made during the pilot phase of a study in Zambia. Despite culturally specific efforts to carefully design and comprehensively adapt, when needed, the study‘s main phase assessments, the participants sometimes failed to respond to assessment stimuli. To reduce potential sources of method bias, the development and piloting stages of these assessments for use in another region in Zambia (Stemler et al., 2009) focused on ensuring a high level of stimulus familiarity. Items and assessment procedures were developed to reflect students‘ low levels of acquaintance with standardized testing and lack of previous test exposure. To increase fairness to examinees, the tasks were designed to be ecologically valid, that is, as similar as possible to the type of tasks that children have to solve in school and everyday life. Still, the otherwise social and friendly students, who were eager to meet visitors and extremely curious about the study‘s activities, became quiet in response to certain stimuli; and the frequency of such quietness was rather alarming.

Zambia-specific efforts have considered the cultural differences impacting assessment design and educational programs, such as the Panga Munthu Test, developed by Robert Serpell and colleagues, in order to avoid biases caused by lack of experience with pencil and paper tests by using a task in which children are instructed to build a man out of clay (Kathuria & Serpell, 1998). Previous reports on assessment development have included similar contexts, namely low-and-middle income countries of sub-Saharan Africa (e.g., Foxcroft, 2011; Grigorenko, 2009; Holding et al., 2004; Tan, Reich, Hart, Thuma, & Grigorenko, 2014), but the unexpected patterns of null responses were not reported. Other

researchers have acknowledged that cultural differences (e.g., students‘ shyness) might lead to a failure to respond to certain assessments (Baddeley, Gardner, & Grantham-McGregor, 1995), but no more specifics were offered. In fact, there is little discussion of the issues on unresponsiveness in cross-cultural literature.

Therefore, we argue that it is important to review missing response patterns and their relation to assessment characteristics (e.g., choice type, stimuli type) as well as relevant cultural differences (e.g., amount of in-class student participation) that could result in unresponsiveness. This report reviews the results of a study in preparation of the Bala Bbala Project (BBP—*Bala Bbala* means ‘read the word‘ in Chitonga, a study of Specific Reading Disabilities, SRD, in rural Southern Province, Zambia) designed to evaluate the main study‘s assessments being used in a new community. After the data were collected, we noticed specific non-random patterns of null responses in the data and therefore engaged in post-hoc analyses to analyze potential factors that might contribute to these patterns. Thus, this report aims to address two questions: (1) How do assessment characteristics (i.e., stimulus type, answer choice, and response type) influence response rates?; and (2) Do age, gender, and socioeconomic status (SES) predict response patterns in the set of items defining the stop rule (i.e., the first 10 items of an assessment)?

Method

Participants

The participants were 207 students (100 female), grades 2–7, from three primary schools, with a mean age of 12.12 years (SD = 2.14 years) who participated in the pilot phase of the BBP (Hein, Reich, Thuma, & Grigorenko, 2014; Tan et al., 2014). The large age range among students in primary schools (7.69 to 17.17 years, mode = 7.86, median = 12.00) is typical of rural Zambia. The reason for this wide age distribution is that in rural Zambia in particular and in sub-Saharan Africa in general, children often take breaks from their studies or repeat grades due to financial difficulties, household responsibilities, health problems, or distance to schools (Mumba, 2002). Despite that, the age-grade correlation was relatively high (Spearman‘s $\rho = .83$, $p < .001$).

Materials

The assessments are a subset of those in the BBP (see Table 1). The following domains were assessed: phonological memory (PM), phonological awareness (PA), expressive vocabulary (EV), mathematics (Zambian Achievement Test-Mathematics, ZAT-M), pseudowords decoding (ZAT-PW), reading comprehension (ZAT-RC), and reading recognition (ZAT-RR). They represent a range of assessment types and their data are coded in a way that is informative to this discussion of missing data. Most of the assessments were used in previous studies in sub-Saharan Africa. EV was adapted from a similar study in Ghana (Grigorenko et al., 2009) and ZAT was adapted from a previous Zambian study in which Zambian curriculum and the University of Zambia was consulted in the development of assessment materials (Stemler et al., 2009). All of the assessments went through a comprehensive adaption process in which items were translated and back-translated by separate native speakers. Items were discussed with the translators, who were from the local

community, with regard to their suitability and relevance. As EV and ZAT were used in similar communities in previous research, adaptations were able to make use of both English and local language versions. PA was developed specifically for use in this community by the translators and linguists with knowledge of the Chitonga phonemic inventory, phonological characteristics, and word formation processes. Since a substantial number of students did not respond to a particular assessment (or at least had a significant amount of missing values), the internal consistencies (Cronbach's α) of the measures used could only be calculated for a smaller number of students (based on a listwise deletion procedure). Most of the internal consistencies were adequate and ranged from .613 for expressive vocabulary to .936 for ZAT-PW. However, the number of students who responded to at least one item ranged from 15 (for expressive vocabulary) to 170 (for ZAT-RR) out of the sample of 207 students. For phonological memory, only 2 students responded to all of the items, which is why Cronbach's α could not be estimated accurately. This observation was one of the reasons that prompted our investigation of missing values in the present study.

Instructions and two practice items per assessment were trialed and validated in the population of the study prior to administering the assessment. All assessors had at least a high school degree but were trained intensively in the data collection protocols. The length of the instructions was held short across all of the assessments and the same instructions were used for several items consecutively to reduce the potential impact of the memory load required by complex instructions on students' responses. Stimuli length was also generally short, but varied in assessments that required longer stimuli (e.g., more difficult phonological memory and reading comprehension items).

Demographic information and socio-economic status (SES)—Age and gender were collected from school records. As part of a larger individually administered questionnaire on the students' development, parents and caretakers responded to five questions about household possessions for indicators of SES (i.e., radio, TV, stove, refrigerator and cell phone). Each of these five questions could be answered with *yes* (= 1) or *no* (= 0).

Phonological Memory (PM)—PM includes 25 items and is designed to assess the students' capacity to memorize the sounds of words with increasing length and difficulty in Chitonga. All stimuli and responses are verbal. For each item, the students are asked to listen to a pseudoword (example: *anza*) and then repeat it.

Phonological Awareness (PA)—PA comprised seven verbal subtests: (1) Initial sound matching (ISM; 10 items; example: word – *meno*, selection for sound matching – *meso, baasa, chukka*); (2) Final sound matching (FSM; 10 items; example: word – *imoota*, selection for sound matching – *igwezo, mwana, mwezi*); (3) Blending syllables (BSyl; 10 items; example: syllables to blend – *ye, bo*, target word – *yebo*); (4) Blending single segments (BSeg; 11 items; example: segments to blend – *a, n, o*, target word – *ano*); (5) Segmenting syllables (SSyl; 11 items; example: word to segment – *kuboko*, target syllables – *ku, bo, ko*); (6) Segmenting single segments (SSeg; 10 items; example: word to segment – *seka*, target segments – *s, e, k, a*); (7) Elision (El; 10 items; example: word – *caano*, sound to

remove – *c*, target word – *aano*). Some subtests were multiple-choice and some were open-ended.

Expressive Vocabulary (EV)—EV required verbal identification of 40 images presented one at a time (see Figure 1, for a sample item).

Zambian Achievement Test (ZAT)—ZAT assesses academic skills and is designed to be administered individually. It was developed based on the curriculum of the Zambian government schools in Chinyanja and English for an earlier study in Eastern Province, Zambia. It was subsequently translated and adapted for use in Chitonga-speaking Southern Province, Zambia for the BBP. The following subtests are included in this report: Mathematics (ZAT-M), Pseudowords (ZAT-PW), Reading Comprehension (ZAT-RC), and Reading Recognition (ZAT-RR). ZAT-M assesses primary school mathematical skills such as number recognition, counting, and multiplication through 60 multiple-choice items with non-verbal responses (pointing). ZAT-PW assesses decoding skills by requiring verbal responses students to read aloud Chitonga pseudowords (example: *pu*) and consists of 40 open-ended items (verbal response, written stimuli). ZAT-RC is a non-verbal task that had a combination of pointing responses to multiple-choice items to assess skills from word-reading through paragraph comprehension (Items 1–18, henceforth labels ZAT-RC-A) and action responses to open-ended items consisting of written directions (Items 19–28, henceforth labels ZAT-RC-B). For each multiple-choice item, the students are asked to read the stimuli and question and then choose their response by pointing to it. The items can be described as single word recognition, following single sentence instructions, short paragraph comprehension.

ZAT-RR assesses the pre-literacy skills of letter recognition/alphabet knowledge and sound-letter correspondence through 40 items that are multiple-choice with non-verbal responses (pointing). For each item, students are instructed to select the most appropriate response to the administrator‘s instructions by pointing to one of four answer choices provided on an answer page. For example, in one item, children are shown a letter and then asked to find the same letter from a set of four letters. Assessment is discontinued after eight consecutive incorrect responses. Figure 1 shows sample items for ZAT-M, ZAT-RR and ZAT-RC.

Procedure

Recruitment was a multi-step process starting with school selection. All schools were within the same district, within 50km of the local hospital that was the center of our study area, and had to have all participating grades (i.e., 3–7). After meeting with head teachers, meetings with parents and caretakers were held before selecting children from each grade and gender combination at random for inclusion. Trained research assistants individually administered assessments at the participants‘ schools and were monitored during data collection. Assessments were given in the local language, Chitonga.

Statistical analysis

Three single factor multivariate analyses of variance (MANOVA) were conducted, examining the effect of stimulus type, answer choice, and response type, respectively, on the

missing responses in the first 10 items (i.e., the set of items defining the stop rule). Contrast coding was used to calculate mean number missing for individuals for each stimulus type, answer choice, and response type for comparison.

A set of logistic regression analyses was carried out to predict the probability of not responding to the first 10 items of an assessment using age, gender, and SES as predictors. Dummy variables (by assessment/subtest; 1 = *no response*, 0 = *some response to at least one of the first 10 items*) were created to indicate whether participants responded to these items. A principal component analysis was conducted using the SES indicators (i.e., radio, TV, stove, refrigerator and cell phone) to extract factor scores for a one-component solution (48.90% of variance explained). The SES factor score was entered as a predictor. Age was entered as a *z*-score.

Results

Descriptive statistics are presented in Table 1. The first 10 items, or all 10 items in the case of subtests with only 10 items, were analyzed because a stop rule of 10 consecutive incorrect responses was used.

Associations between response rates and characteristics of an assessment

First, there were significant differences in the mean number of missing values between the different stimuli types [$F(3, 203) = 110.97$, $p = .000$; Wilk's $\lambda = 0.38$; partial $\eta^2 = .62$]. Missing values were highest for assessments containing written stimuli ($M = 5.79$, $SD = 4.35$; 3 subtests), lower for verbal ($M = 2.74$, $SD = 3.27$; 8 subtests), lower for visual ($M = 1.46$, $SD = 2.75$; 1 subtest), and lowest for items that contained both written and visual material ($M = 0.31$, $SD = 1.24$; 2 subtests). Second, there were significant differences in the number of missing values between open and closed answer choices [$F(1, 206) = 108.25$, $p = .000$; Wilk's $\lambda = 0.66$; partial $\eta^2 = .34$]. Missing values were higher for assessments requiring open-ended answers ($M = 3.24$, $SD = 2.72$; 9 subtests) compared to closed answers ($M = 2.01$, $SD = 1.88$; 5 subtests). Finally, there were significant differences in the number of missing values between the three required response types [$F(2, 134) = 71.58$, $p = .000$; Wilk's $\lambda = 0.48$; partial $\eta^2 = .52$]. Missing values were highest for assessments that require action responses ($M = 4.86$, $SD = 4.55$; only items 19 to 28 of ZAT-RC required an action response), lower for assessments that require verbal responses ($M = 2.56$, $SD = 2.97$; 11 subtests required a verbal response) and lowest for assessments that allow either pointing or verbal responses ($M = 1.34$, $SD = 1.88$; 2 subtests).

Prediction of probability of non-responding

The χ^2-values of the omnibus test of all coefficients in the model showed that the three predictors did not significantly contribute to the prediction of being a non-responder for most of the assessments (range: 0.149 for PA-ISM to 7.554 for PA-El). However, significant χ^2-values and significant partial contributions were found for ZAT-PW, ZAT-RC-A and ZAT-RC-B.

For ZAT-PW the full model omnibus test was significant [χ^2 (3) = 28.580, $p < .001$], yielding an overall prediction success of 67% (77.7% for responders and 52.6% for non-

responders), with a small relationship between the predictor set and the predicted grouping (Nagelkerke's R^2 = .198). Age (B = −.802, $SE\ B$ = .179, p < .001; Exp(B) = .449) and SES (B = −.517, $SE\ B$ = .176, p = .003; Exp(B) = .596) made significant partial contributions, but not gender. The odds ratio values indicated that every 1 SD increase in age and SES are associated with a 55.1% and 40.4% decrease, respectively, in the odds of not responding. For ZAT-RC-A, the model omnibus test was significant [χ^2 (3) = 61.093, p < .001]. Prediction success overall was 73.7% (79.6% for responders, 66.7% for non-responders), with a medium relationship between the predictor set and the predicted grouping (Nagelkerke's R^2 = .387). Age (B = −1.280, $SE\ B$ = .217, p < .001; Exp(B) = .278) and SES (B = −.894, $SE\ B$ = .206, p < .001; Exp(B) = .409) made significant partial contributions. The odds ratio values indicated that every 1 SD increase in age and SES are associated with a 72.2% and 59.1% decrease, respectively, in the odds of not responding on ZAT-RC-A. A similar pattern of results was obtained for ZAT-RC-B, where the full model omnibus test was significant [χ^2 (3) = 12.001, p = .005]. However, prediction success overall was 77.1%, with 99.3% for responders but only 0% for non-responders, with a small relationship between the predictor set and the predicted grouping (Nagelkerke's R^2 = .107). Only age (B = −0.689, $SE\ B$ = .203, p < .001; Exp(B) = .502) made a significant partial contribution.

Discussion

Stimulus type, answer choice, and response type differentially impact responsiveness, and, thus, missing value patterns. Missing values were highest for written stimuli, open-ended answer choices, and action responses, entailing methodological and conceptual implications. All of these item features are things that put the children in unfamiliar situations, namely those that require them to respond through action – oral or non-oral – rather than through passive pointing. Everyday lessons in rural Zambian classrooms are highly teacher oriented, and an average student-teacher ratio of about 49 (UNESCO Institute for Statistics, 2014) leaves only little room for student-teacher and peer interactions and individualized learning. Thus, it is particularly important to pay close attention to the ecological validity of assessments so that they best reflect the types of response that are required from students in their everyday classroom interactions.

Methodologically, these findings highlight that analyzing patterns of participants' responsiveness is a necessary precursor to investigating linguistic, cultural, functional and metric equivalence. For instance, calculating the percentage of correctly solved items as an indicator of item difficulty is only possible if participants respond to items. The significant associations between age and SES and the missing responses makes it necessary to consider appropriate methods to treat data that are not missing at random (cf., Schafer & Graham, 2002). Future studies should aim at investigating the potential pathways through which factors such as age and SES may contribute to missing values on certain assessments. In the present study, age and SES did partially contribute to explaining the probability of non-responses on the assessments of pseudoword reading and reading comprehension, but not with regard to other assessments. Both skills are rather advanced skills and are most likely related to educational experiences in school. Thus, older students might attempt solving these items while younger students might not had the opportunities to practice skills pertaining to the decoding of pseudowords or comprehension of written text,

and could thus decide not to answer an item. Moreover, the SES indicators in this study assess the availability of household items and do therefore reflect the level of economic wealth of a particular household. The finding that higher SES is associated with a lower likelihood of non-responding it indicative of differences in the experiences that the children have by SES that might contribute to confidence, assertiveness, or the willingness to take chances.

Methods of cross-cultural and cross-linguistic translation, adaptation, and design are improving with new guidelines and the work of the ITC and other research groups. Yet, the results illustrate that assessment piloting and consideration of cultural specifics are imperative. Even with advanced examination techniques for assessment development (e.g., differential item functioning; Hambleton & Kanjee, 1995), we should not underestimate the impact of stimulus types and response types during the adaption process of an assessment. Further, we should not assume that variation in response patterns would be the same for participants of similar linguistic (e.g., another language from the group of Bantu languages, to which Chitonga belongs) or geographic backgrounds (e.g., another sub-Saharan contexts).

Conceptually, a discussion of the role of assessment characteristics in differential response patterns is evident, for example, in recent efforts from sub-Saharan Africa (Alcock, Holding, Mung'ala-Odera, & Newton, 2008; Solarsh & Alant, 2006). The accumulating knowledge about the role of assessment characteristics might lead to higher-level project changes rather than individual assessment changes. For example, instead of developing a new assessment battery, the BBP continued using the assessments as designed, but later in the study, when missing data would not impact screening. Further, the data collectors, who were all from the local community, were consulted and observations of the children during data collections were completed, leading to training of data collectors to be more encouraging of responsiveness as well as to probe the children on the nature of their continued unresponsiveness. More specifically, response choices of "Child Doesn't Know" and "No Response Given" were added. There are instances in which some difficult items, or items likely to result in unresponsiveness might be informative. However, it is obvious that assessments resulting in large amounts of missing data should not be used in screening measures. Further, future studies may want to consider using narrower age and grade ranges. The expectation of this study was that younger children would not perform at the same level as the older children; however, it was expected that they would answer more items incorrectly, not that they would be more unresponsive.

The present study emphasizes that the development of measures of reading-related and other cognitive skills has to be tailored to the experiences of the students in this community, including experiences at school, at home, and elsewhere in the community. For example, the assessment of pseudoword reading (the assessment with the highest percentage of non-responders) in this study required the child to read pseudowords aloud, thus entailing a mode of open-ended, verbal responses to written stimuli that may not be familiar to each student. As digitalization and computerization of assessment continues to be on the rise, it is likely that these assessments will reach even remote communities in sub-Saharan Africa. An assessment of pseudoword decoding could be altered in a way that requires the student to "point and click" on four alternatives of recorded pseudowords to demonstrate their skill of

decoding a pseudoword, which may lead to fewer rates of missing values on an assessment. Thus, we recommend that different response types (e.g., verbal, action) should be thoroughly piloted to examine the effect of assessment characteristics on students‘ responses. This work would benefit from a mixed methods approach that includes qualitative or ethnographic studies to capture the spectrum of skills relevant to the development of cognitive skills in these communities, as well as the repertoire of educational experiences. There are a number of examples in the literature that have used a qualitative approach to establish criterion validity of the constructs under study (e.g., the assessment of depression-like and anxiety-like problems of war-affected adolescents in northern Uganda; Betancourt et al., 2009), and that can inform the integration of qualitative methods into the study of patterns of non-responses. In light of an increase in attention to the development of cross-cultural (van de Vijver, 2002) and multicultural assessments (Dana, 2005; Grigorenko, 2009), we hope that this illustration of an unexpected cultural difference will contribute to discussions of how to address it.

Acknowledgments

This report was supported by NIH R01 TW008274. This report does not necessarily represent the positions of the NIH as grantees are encouraged to express their professional judgment. We thank Mei Tan and our colleagues at Macha Research Trust for their contributions.

References

Alcock KJ, Holding PA, Mung’ala-Odera V, Newton CRJC. Constructing tests of cognitive abilities for schooled and unschooled children. Journal of Cross-Cultural Psychology. 2008; 39(5):529–551.10.1177/0022022108321176

Baddeley A, Gardner JM, Grantham-McGregor S. Cross-cultural cognition: developing tests for developing countries. Applied Cognitive Psychology. 1995; 9:S173–S195.

Betancourt TS, Bass J, Borisova I, Neugebauer R, Speelman L, Onyango G, Bolton P. Assessing local instrument reliability and validity: a field-based example from northern Uganda. Social Psychiatry and Psychiatric Epidemiology. 2009; 44(8):685–692.10.1007/s00127-008-0475-1 [PubMed: 19165403]

Dana, RH., editor. Multicultural assessment: principles, applications, and examples. Mahwah, NJ: Lawrence Erlbaum Associates; 2005.

Foxcroft CD. Ethical issues related to psychological testing in Africa: What I have learned (so far). Online Readings in Psychology and Culture. 2011; 2(2) doi:10.9707/.

Grigorenko, EL., editor. Multicultural psychoeducational assessment. New York, NY: Springer; 2009.

Grigorenko, EL.; Nedwick, K.; Kwadade, D.; Boro, E.; Hart, L.; Newman, T.; Jarvin, L. Assessing mother tongue in the era of globalization: Primise and challange. In: Grigorenko, EL., editor. Multicultural psychoeducational assessment. New York, NY: Springer; 2009. p. 187-212.

Hambleton RK, Kanjee A. Increasing the validity of cross-cultural assessments: use of improved methods for test adaptations. European Journal of Psychological Assessment. 1995; 11(3):147–157.

Hein S, Reich J, Thuma PE, Grigorenko EL. Physical growth and non-verbal intelligence: Associations in Zambia. The Journal of Pediatrics. 2014; 165(5):1017–1023. e1011. [PubMed: 25217196]

Holding PA, Taylor HG, Kazungu SD, Mkala T, Gona J, Mwamuye B, Stevenson J. Assessing cognitive outcomes in a rural African population: Development of a neuropsychological battery in Kilifi District, Kenya. Journal of the International Neuropsychological Society. 2004; 10:246–260. [PubMed: 15012845]

International Test Commission. The ITC 5th International Conference on Psychological and Educational Test Adaptation across Language and Cultures; 2013; 2006. from http://www.intestcom.org/itc_conference_2006.htm

Kathuria R, Serpell R. Standardization of the Panga Munthu Test-A nonverbal cognitive test developed in Zambia. The Journal of Negro Education. 1998; 67(3):228–241.

Mumba, EC. Education for all: Increasing access to education for girls in Zambia. Paper presented at the 2nd Pan-Commonwealth Forum on Open Learning; Durban, South Africa. 2002.

Peña ED. Lost in translation: methodological considerations in cross-cultural research. Child Development. 2007; 78(4):1255–1264. [PubMed: 17650137]

Schafer JL, Graham JW. Missing data: Our view of the state of the art. Psychological Methods. 2002; 7(2):147–177. [PubMed: 12090408]

Solarsh B, Alant E. The challenge of cross-cultural assessment--The Test of Ability To Explain for Zulu-speaking Children. Journal of Communication Disorders. 2006; 39(2):109–138.10.1016/j.jcomdis.2005.11.002 [PubMed: 16434049]

Stemler, SE.; Chamvu, F.; Chart, H.; Jarvin, L.; Jere, J.; Hart, L.; Grigorenko, EL. Assessing competencies in reading and mathematics in Zambian children. In: Grigorenko, EL., editor. Multicultural psychoeducational assessment. New York, NY: Springer; 2009. p. 157-186.

Tan M, Reich J, Hart L, Thuma PE, Grigorenko EL. Examining the specific effects of context on adaptive behavior and achievement in a rural African community: Six case studies from rural areas of Southern Province, Zambia. Journal of Autism and Developmental Disorders. 2014; 44(2):271–282.10.1007/s10803-012-1487-y [PubMed: 22391811]

UNESCO Institute for Statistics. Primary education (ISCED 1) Pupil-teacher ratio. 2014. Retrieved June 1, 2014, from http://data.un.org/Data.aspx?d=UNESCO&f=series%3APTRHC_1

van de Vijver FJR. Cross-cultural assessment: value for money? Applied Psychology. 2002; 51(4):545–566.

Van de Vijver FJR, Hambleton RK. Translating tests: some practical guidelines. European Psychologist. 1996; 1(2):89–99.

van de Vijver FJR, Poortinga YH. Towards an integrated analysis of bias in cross-cultural assessment. European Journal of Psychological Assessment. 1997; 13(1):29–37.

van de Vijver, FJR.; Poortinga, YH. Conceptual and methodological issues in adapting tests. In: Hambleton, RK.; Merenda, PF.; Spielberger, CD., editors. Adapting educational and psychological tests for cross-cultural assessment. Mahwah, NJ: Lawrence Erlbaum; 2005. p. 39-63.

Highlights

- Assessment characteristics differentially impact patterns of responsiveness
- Using written stimuli resulted in more missing values compared to visual stimuli
- Open-ended answer choices generated more missing values than closed answer choices
- Actions responses resulted in more missing values than other responses
- Age and socio-economic status accounted for some variation in missing values

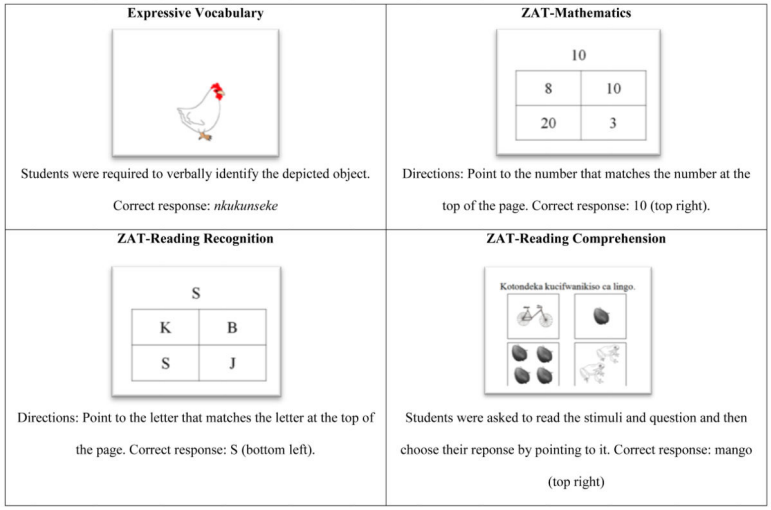

Figure 1.
Sample items for expressive vocabulary and three subtests of the Zambian Achievement Test (ZAT).

Table 1

Description and descriptive statistics of assessments

Assessment	# Items	Stimuli	Choice	Response	Non-responders (%)	All responders	Mixed	$M_{\text{Number missing values}}$
PA-BSyb	10	Verbal	Open	Verbal	62 (30.0)	76	66	4.59
PA-BSeg	10	Verbal	Open	Verbal	41 (19.8)	153	12	2.19
PA-El	10	Verbal	Open	Verbal	52 (25.1)	107	45	3.55
PA-SSyl	10	Verbal	Open	Verbal	64 (30.9)	97	43	4.24
PA-SSeg	11	Verbal	Open	Verbal	43 (20.8)	144	17	2.46
PA-ISM	10	Verbal	Closed	Verbal	37 (17.9)	162	3	1.81
PA-FSM	10	Verbal	Closed	Verbal	40 (19.3)	149	17	2.23
PM	20	Verbal	Open	Verbal	11 (5.3)	177	19	0.70
EV	40	Visual	Open	Verbal	17 (8.2)	108	82	1.46
ZAT-RC-A	18	Writ/Vis	Closed	Point/Verb	95 (45.9)	83	29	5.33
ZAT-RC-B	10	Written	Open	Action	47 (22.7)	50	39	3.19
ZAT-PW	40	Writ/Vis	Open	Verbal	89 (43.0)	79	34	5.20
ZAT-RR	40	Written	Closed	Point/Verb	7 (3.4)	199	1	0.25
ZAT-M	60	Written	Closed	Verbal	7 (3.4)	187	13	0.37

Notes. PA = phonological awareness. BSyb = blending syllables. BSeg = blending single segments. El = elision. SSyl = segmenting syllables. SSeg = segmenting single segments. ISM = initial sound matching. FSM = final sound matching. PM = phonological memory. EV = expressive vocabulary. ZAT = Zambian Achievement Test. RC = reading comprehension. PW = pseudoword reading. M = mathematics. RR = reading recognition. Types of stimuli were coded as 1 (*verbal*), 2 (*written*), 3 (*written and visual*), and 4 (*visual*). Answer choices were coded as 1 (*closed answer*) and 2 (*open answer*). Types of responses were coded as 1 (*verbal*), 2 (*pointing or verbal*), and 3 (*action*). ZAT-RC uses two different assessment modalities. Items 1 to 18 use written and visual stimuli, require closed answers using either a pointing or verbal response. Items 19 to 28 use written stimuli only and allow open responses using an action response. Due to the stop rule, items 19 to 28 were not administered to some of the students.